

S. Zhou

Nerun

1631

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/549,827A

TIME: 10:54:14

Input Set : A:\09549827supplsequencelisting.txt

Output Set: N:\CRF3\12202001\I549827A.raw

ENTERED

3 <110> APPLICANT: Rzhetsky, Andrey
 4 Kalachikov, Sergey
 5 Krauthammer, Michael
 6 Friedman, Carol
 7 Kra, Pauline
 9 <120> TITLE OF INVENTION: GENE DISCOVERY THROUGH COMPARISONS OF
 10 NETWORKS OF STRUCTURAL AND FUNCTIONAL RELATIONSHIPS AMONG
 11 GENES AND PROTEINS
 14 <130> FILE REFERENCE: A31869-A 070050.1046
 C--> 16 <140> CURRENT APPLICATION NUMBER: US/09/549,827A
 17 <141> CURRENT FILING DATE: 2000-04-14
 19 <160> NUMBER OF SEQ ID NOS: 22
 21 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 39
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Artificial Sequence
 28 <220> FEATURE:
 29 <223> OTHER INFORMATION: Prophetic example of coded message
 31 <400> SEQUENCE: 1
 32 agcaactaaa caccatcca agcaaacaca cacacaaac 39
 34 <210> SEQ ID NO: 2
 35 <211> LENGTH: 40
 36 <212> TYPE: DNA
 37 <213> ORGANISM: Artificial Sequence
 39 <220> FEATURE:
 40 <223> OTHER INFORMATION: Prophetic example of coded message
 42 <400> SEQUENCE: 2
 43 aagcaactaa acaccatcc aagcaaacac acacacaaac 40
 45 <210> SEQ ID NO: 3
 46 <211> LENGTH: 292
 47 <212> TYPE: DNA
 48 <213> ORGANISM: Artificial Sequence
 50 <220> FEATURE:
 51 <223> OTHER INFORMATION: Prophetic example of coded message
 53 <400> SEQUENCE: 3
 54 aagtagagat ccacggaagg aacgatccaa acaaagacgc aacgacagaa ataacgatcc 60
 55 acataactat ccaaatacat acgcacggaa gtacacacgt aattaaacac ggaagtacat 120
 56 acagatccat ccacggatcc aaataacgaa ttaattacgc atccaaacaa atacggaagt 180
 57 actcaaacac ggaacgaacc atccacggaa ggacctacat acgtaagcaa ggatccacgg 240
 58 aaggaacgaa gtacctatcc aaacacagac ggaagtaagc aacgacagat cc 292
 60 <210> SEQ ID NO: 4
 61 <211> LENGTH: 10
 62 <212> TYPE: DNA
 63 <213> ORGANISM: Artificial Sequence
 65 <220> FEATURE:
 66 <223> OTHER INFORMATION: Prophetic example of coded message

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/549,827A

TIME: 10:54:14

Input Set : A:\09549827supplsequencelisting.txt

Output Set: N:\CRF3\12202001\I549827A.raw

```

68 <400> SEQUENCE: 4
69 atctgtcacg 10
71 <210> SEQ ID NO: 5
72 <211> LENGTH: 405
73 <212> TYPE: DNA
74 <213> ORGANISM: Human
76 <400> SEQUENCE: 5
77 catggcttcc tggacaccaa ccctgccatc cgggagcaga cgggtcaagtc catgctgctc 60
78 ctggccccaag agctgaacga ggccaacctc aatgtggagc tgatgaagca ctttgcacgg 120
79 ctacaggcca aggatgaaca gggcccatc cgtgcaaca ccacagtctg cctgggcaaa 180
80 atcggctcct acctcagtgc tagcaccaga cacagggtcc ttacctctgc cttcagccga 240
81 gccactaggg acccgtttgc accgtcccg gttgcgggtg tcttgggctt tgctgccacc 300
82 cacaacctct actcaatgaa cgactgtgcc cagaagatcc tgcctgtgct ctgcggtctc 360
83 actytagatc ctgagaaatc cgtgcgagac caggccttca aggca 405
85 <210> SEQ ID NO: 6
86 <211> LENGTH: 453
87 <212> TYPE: DNA
88 <213> ORGANISM: Human
90 <220> FEATURE:
91 <221> NAME/KEY: variation
92 <222> LOCATION: (146)...(146)
93 <223> OTHER INFORMATION: A, C, G, or T
95 <400> SEQUENCE: 6
96 ccttcgagtt cggcaatgct ggggcccgtt tcttcacgcc cctcttcaag gtgggcaagt 60
97 tcttgagcgc tgaggagtat cagcagaaga tcatccctgt ggtggtcaag atgttctcat 120
W--> 98 ccaactgacc ggccatgcgc atccgnetcc tgcagcagat ggagcagttc atccagtacc 180
99 ttgacgagcc aacagtcaac acccagatct tccccacgt cgtacatggc ttcttgga 240
100 ccaaccctgc catccgggag cagaagggtc agtccatgct gctcctggcc ccaaagctga 300
101 acgaggccaa cctcaatgtg gagctgatga agcatttgc acggctacag gccaaggatg 360
102 aacaggggccc catccgtgc aacaccacaa cctgcctggg caaatcggc tctacctca 420
103 gtgctagcac cagacacagg gtccttacct ctg 453
105 <210> SEQ ID NO: 7
106 <211> LENGTH: 1727
107 <212> TYPE: DNA
108 <213> ORGANISM: Human
110 <400> SEQUENCE: 7
111 cagccgaagc amgcaaaaat tcttccagga gctgagcaag agcctggacg cattccctga 60
112 ggayttctgt cggcacaagg tgctgcccc aatgtgtgac gccttcgagt tcggcaatgc 120
113 tggggccggt gtcttcacgc ccctcttcaa ggtgggcaag ttcttgagcg ctgaggagta 180
114 tcagcagaag atcatccctg tgggtgtcaa gatgttctca tccactgacc gggccatgcg 240
115 catccgcctc ctgcagcaga tggagcagtt catccagtac cttgacgagc caacagtcaa 300
116 caccagatc tccccccacg tcgtacatgg cttcctggac accaaccctg ccatccggga 360
117 gcagacggtc aagtccatgc tgctcctggc ccaaagctg aacgaggcca acctcaatgt 420
118 ggagctgatg aagcatttgc cagggctaca ggccaaggat gaacagggcc ccatccgtg 480
119 caacaccaca gtctgcctgg gcaaaatcgg ctcctacctc agtgctagca ccagacacag 540
120 ggtccttacc tctgccttca gccgagccac tagggaccg tttgcaccgt cccgggttgc 600
121 ggggttctct ggtcttctgc ccaccacaa cctctactca atgaacgact gtgccagaa 660
122 gatcctgcct gtgctctgcg gtctcactgt agatcctgag aaatccgtgc gagaccaggc 720
123 cttcaaggcm wtccggagct tctgttccaa attggagtct gtgtcggagg acccgacca 780

```

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/549,827A

TIME: 10:54:14

Input Set : A:\09549827supplsequencelisting.txt

Output Set: N:\CRF3\12202001\I549827A.raw

```

124 gctggaggaa gtggagaagg atgtccatgc agcctccagc cctggcatgg gaggagccgc 840
125 agctagctgg gcaggctggg cgtgaccggg gtctctcac tcacctcaa gctgatccgt 900
126 tcgcacccaa cactgcccc aacagaaacc aacattcccc aaagaccac gcctgaagga 960
127 gttcctgccc cagccccac ccctgttctt gccacccta caacctcagg ccactgggag 1020
128 acgcaggagg aggacaagga cacagcagag gacagcagca ctgctgacag atgggacgac 1080
129 gaagactggg gcagcctgga gcaggaggcc gagtctgtgc tggcccagca ggacgactgg 1140
130 agcaccgggg gccaaagtga cgtgctagt caggtcagca actccgacca caaatcctcc 1200
131 aaatccccag agtccgactg gagcagctgg gaartcagg gctcctggga acagggctgg 1260
132 caggagccaa gctcccagga gccacctyct gacggtacac ggctggccag cgagtataac 1320
133 tggggtggcc cagagtccag cgacaagggc gacccttctg ctacctgtc tgcacgtccc 1380
134 agcaccagc cgaggccaga ctcttggggg gaggacaact gggagggcct cgagactgac 1440
135 agtcgacagg tcaaggctga gctggccggg aagaagcgcg aggagcggcg gcgggagatg 1500
136 gaggccaaac gcgccgagag gaaggtgcc aaggcccat gaagctggga gcccggaagc 1560
137 tggactgaac cgtggcgggt gcccttccg gctgcggaga gcccgcccca cagatgtatt 1620
138 tattgtacaa accatgtgag cccggccggc cagccaggcc atctcacgtg tacataatca 1680
139 gagccacaat aaattctatt tcacaaaaaa aaaaaaaaa aaaaaaa 1727

```

141 <210> SEQ ID NO: 8

142 <211> LENGTH: 287

143 <212> TYPE: PRT

144 <213> ORGANISM: Human

146 <220> FEATURE:

147 <221> NAME/KEY: VARIANT

148 <222> LOCATION: (4)...(4)

149 <223> OTHER INFORMATION: Any amino acid

151 <221> NAME/KEY: VARIANT

152 <222> LOCATION: (244)...(244)

153 <223> OTHER INFORMATION: Any amino acid

155 <400> SEQUENCE: 8

```

W--> 156 Ser Arg Ser Xaa Gln Lys Phe Phe Gln Glu Leu Ser Lys Ser Leu Asp
157 1 5 10 15
158 Ala Phe Pro Glu Asp Phe Cys Arg His Lys Val Leu Pro Gln Leu Leu
159 20 25 30
160 Thr Ala Phe Glu Phe Gly Asn Ala Gly Ala Val Val Leu Thr Pro Leu
161 35 40 45
162 Phe Lys Val Gly Lys Phe Leu Ser Ala Glu Glu Tyr Gln Gln Lys Ile
163 50 55 60
164 Ile Pro Val Val Val Lys Met Phe Ser Ser Thr Asp Arg Ala Met Arg
165 65 70 75 80
166 Ile Arg Leu Leu Gln Gln Met Glu Gln Phe Ile Gln Tyr Leu Asp Glu
167 85 90 95
168 Pro Thr Val Asn Thr Gln Ile Phe Pro His Val Val His Gly Phe Leu
169 100 105 110
170 Asp Thr Asn Pro Ala Ile Arg Glu Gln Thr Val Lys Ser Met Leu Leu
171 115 120 125
172 Leu Ala Pro Lys Leu Asn Glu Ala Asn Leu Asn Val Glu Leu Met Lys
173 130 135 140
174 His Phe Ala Arg Leu Gln Ala Lys Asp Glu Gln Gly Pro Ile Arg Cys
175 145 150 155 160
176 Asn Thr Thr Val Cys Leu Gly Lys Ile Gly Ser Tyr Leu Ser Ala Ser

```

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/549,827A

TIME: 10:54:14

Input Set : A:\09549827supplsequencelisting.txt

Output Set: N:\CRF3\12202001\I549827A.raw

```

177                               165                               170                               175
178 Thr Arg His Arg Val Leu Thr Ser Ala Phe Ser Arg Ala Thr Arg Asp
179                               180                               185                               190
180 Pro Phe Ala Pro Ser Arg Val Ala Gly Val Leu Gly Phe Ala Ala Thr
181                               195                               200                               205
182 His Asn Leu Tyr Ser Met Asn Asp Cys Ala Gln Lys Ile Leu Pro Val
183                               210                               215                               220
184 Leu Cys Gly Leu Thr Val Asp Pro Glu Lys Ser Val Arg Asp Gln Ala
185 225                               230                               235                               240
W--> 186 Phe Lys Ala Xaa Arg Ser Phe Leu Ser Lys Leu Glu Ser Val Ser Glu
187                               245                               250                               255
188 Asp Pro Thr Gln Leu Glu Glu Val Glu Lys Asp Val His Ala Ala Ser
189                               260                               265                               270
190 Ser Pro Gly Met Gly Gly Ala Ala Ala Ser Trp Ala Gly Trp Ala
191                               275                               280                               285
194 <210> SEQ ID NO: 9
195 <211> LENGTH: 223
196 <212> TYPE: PRT
197 <213> ORGANISM: Human
199 <400> SEQUENCE: 9
200 Val Met Glu Leu Leu Glu Glu Asp Leu Thr Cys Pro Ile Cys Cys Ser
201 1                               5                               10                               15
202 Leu Phe Asp Asp Pro Arg Val Leu Pro Cys Ser His Asn Phe Cys Lys
203                               20                               25                               30
204 Lys Cys Leu Glu Gly Ile Leu Glu Gly Ser Val Arg Asn Ser Met Trp
205                               35                               40                               45
206 Arg Pro Ala Pro Phe Lys Cys Pro Thr Cys Arg Lys Glu Thr Ser Ala
207 50                               55                               60
208 Thr Gly Ile Asn Ser Leu Gln Val Asn Tyr Ser Leu Lys Gly Ile Val
209 65                               70                               75                               80
210 Glu Lys Tyr Asn Lys Ile Lys Ile Ser Pro Lys Met Pro Val Cys Lys
211                               85                               90                               95
212 Gly His Met Gly Gln Pro Leu Asn Ile Phe Cys Leu Thr Asp Met Gln
213                               100                              105                              110
214 Leu Ile Cys Gly Ile Cys Ala Thr Arg Gly Glu His Thr Lys His Val
215                               115                              120                              125
216 Phe Cys Ser Ile Glu Asp Ala Tyr Ala Gln Glu Arg Asp Ala Phe Glu
217                               130                              135                              140
218 Ser Leu Phe Gln Ser Phe Glu Thr Trp Arg Arg Gly Asp Ala Leu Ser
219 145                              150                              155                              160
220 Arg Leu Asp Thr Met Glu Thr Ser Lys Arg Lys Ser Leu Gln Leu Met
221                               165                              170                              175
222 Thr Lys Asp Ser Asp Lys Val Lys Glu Phe Phe Glu Lys Leu Gln His
223                               180                              185                              190
224 Thr Leu Asp Gln Lys Lys Asn Glu Ile Leu Ser Asp Phe Glu Thr Met
225                               195                              200                              205
226 Lys Leu Ala Val Met Gln Ala Tyr Asp Pro Glu Ile Asn Lys Leu
227                               210                              215                              220
230 <210> SEQ ID NO: 10

```

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/09/549,827A

TIME: 10:54:14

Input Set : A:\09549827supplsequencelisting.txt

Output Set: N:\CRF3\12202001\I549827A.raw

```

231 <211> LENGTH: 218
232 <212> TYPE: PRT
233 <213> ORGANISM: Mouse
235 <400> SEQUENCE: 10
236 Val Leu Glu Met Ile Lys Glu Glu Val Thr Cys Pro Ile Cys Leu Glu
237 1 5 10 15
238 Leu Leu Lys Glu Pro Val Ser Ala Asp Cys Asn His Ser Phe Cys Arg
239 20 25 30
240 Ala Cys Ile Thr Leu Asn Tyr Glu Ser Asn Arg Asn Thr Asp Gly Lys
241 35 40 45
242 Gly Asn Cys Pro Val Cys Arg Val Pro Tyr Pro Phe Gly Asn Leu Arg
243 50 55 60
244 Pro Asn Leu His Val Ala Asn Ile Val Glu Arg Leu Lys Gly Phe Lys
245 65 70 75 80
246 Ser Ile Pro Glu Glu Glu Gln Lys Val Asn Ile Cys Ala Gln His Gly
247 85 90 95
248 Glu Lys Leu Arg Leu Phe Cys Arg Lys Asp Met Met Val Ile Cys Trp
249 100 105 110
250 Leu Cys Glu Arg Ser Gln Glu His Arg Gly His Gln Thr Ala Leu Ile
251 115 120 125
252 Glu Glu Val Asp Gln Glu Tyr Lys Glu Lys Leu Gln Gly Ala Leu Trp
253 130 135 140
254 Lys Leu Met Lys Lys Ala Lys Ile Cys Asp Glu Trp Gln Asp Asp Leu
255 145 150 155 160
256 Gln Leu Gln Arg Val Asp Trp Glu Asn Gln Ile Gln Ile Asn Val Glu
257 165 170 175
258 Asn Val Gln Arg Gln Phe Lys Gly Leu Arg Asp Leu Leu Asp Ser Lys
259 180 185 190
260 Glu Asn Glu Glu Leu Gln Lys Leu Lys Lys Glu Lys Lys Glu Val Met
261 195 200 205
262 Glu Lys Leu Glu Glu Ser Glu Asn Glu Leu
263 210 215
266 <210> SEQ ID NO: 11
267 <211> LENGTH: 124
268 <212> TYPE: PRT
269 <213> ORGANISM: Mouse
271 <400> SEQUENCE: 11
272 Met Glu Pro Val Ala Ser Asn Ile Gln Val Leu Leu Gln Ala Ala Glu
273 1 5 10 15
274 Phe Leu Glu Arg Arg Glu Arg Glu Ala Glu His Gly Tyr Ala Ser Leu
275 20 25 30
276 Cys Pro His His Ser Pro Gly Thr Val Cys Arg Arg Arg Lys Pro Pro
277 35 40 45
278 Leu Gln Ala Pro Gly Ala Leu Asn Ser Gly Arg Ser Val His Asn Glu
279 50 55 60
280 Leu Glu Lys Arg Arg Arg Ala Gln Leu Lys Arg Cys Leu Glu Gln Leu
281 65 70 75 80
282 Arg Gln Gln Met Pro Leu Gly Val Asp Cys Thr Arg Tyr Thr Thr Leu
283 85 90 95

```

Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

DATE: 12/20/2001

PATENT APPLICATION: US/09/549,827A

TIME: 10:54:15

Input Set : A:\09549827supplsequencelisting.txt

Output Set: N:\CRF3\12202001\I549827A.raw

L:16 M:270 C: Current Application Number differs, Replaced Current Application Number
L:98 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6
L:156 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:186 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:341 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:381 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:382 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:385 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:386 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:593 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:632 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22